

# **ARL PARA O GRÁFICO DE CONTROLE DA MEDIANA QUANDO O ESTIMADOR DO DESVIO PADRÃO É BASEADO NOS QUARTIS AMOSTRAIS**

## **Felipe Martins**

Departamento de Estatística, CCEN - UFPA.  
Rua Augusto Corrêa, 01 - Campus Universitário do Guamá.  
CEP 66075-110 Belém-Pará-Brasil  
felipe@ufpa.br

## **Jaciane do Carmo Ribeiro**

Departamento de Estatística, CCEN - UFPA.  
Rua Augusto Corrêa, 01 - Campus Universitário do Guamá.  
CEP 66075-110 Belém-Pará-Brasil  
jaciane@ufpa.br

## **Silvia dos Santos de Almeida**

Departamento de Estatística, CCEN - UFPA.  
Rua Augusto Corrêa, 01 - Campus Universitário do Guamá.  
CEP 66075-110 Belém-Pará-Brasil  
salmeida@ufpa.br

## **Edson Marcos Leal Soares Ramos**

Departamento de Estatística, CCEN - UFPA.  
Rua Augusto Corrêa, 01 - Campus Universitário do Guamá.  
CEP 66075-110 Belém-Pará-Brasil  
edson@ufpa.br

## **Resumo**

Todo e qualquer serviço prestado ou item produzido possui uma certa variabilidade na característica de interesse (variável), sendo que quanto menor a variabilidade gerada pela característica em questão, maior a qualidade e o valor agregado a este bem. O monitoramento constante da variabilidade e a busca de maneiras para reduzi-la são fatores importantes na obtenção de um produto de boa qualidade. Neste contexto, a utilização das técnicas do Controle Estatístico da Qualidade (*CEQ*) é indispensável, onde se destaca, no controle da variabilidade, a utilização dos gráficos de controle, os quais podem detectar causas identificáveis de variação. Os gráficos de controle testam a hipótese de que o processo encontra-se sob controle estatístico. Entretanto, ao realizar algum teste de hipótese é possível tomar a decisão errada. Com o gráfico de controle não é diferente, podendo-se cometer dois erros, o Erro Tipo I e o Erro Tipo II. O primeiro supõe que o processo está fora de controle estatístico, quando na verdade ele está sob controle e o segundo supõe que o processo está sob controle estatístico, quando na verdade ele se encontra fora de controle estatístico (Ramos *et al.*, 2005). Assim, é necessário controlar estes erros e através da estatística do *ARL* (*Average Run Length*) pode-se avaliar quão bom é o gráfico de controle em relação à detecção desses erros, onde o *ARL* determina o número médio de amostras necessárias até obter um sinal de que o processo encontra-se fora de controle estatístico (Erro Tipo I e Erro Tipo II). Desta forma, este trabalho propõe uma nova metodologia para elaborar o *ARL* para o gráfico de controle da mediana, onde este é construído utilizando-se um estimador do desvio padrão populacional baseado nos quartis amostrais.

**Palavras-Chaves:** *ARL*; Quartis; Gráficos de Controle.

### Abstract

Every and each service that is done or item that is produced has a kind of variability in the interest characteristic (variable), but in the truth as less variability created by the characteristic talked before, bigger are the quality and the lodger value to this item. The constant monitoring of the variability and the search of different ways to reduce it are important factors to obtain a product with good quality. In this context, the use of statistical quality control (*SQC*) techniques is indispensable where it shows the variability control, the use of control charts, where it can be detected when there is some unknown variation cause. The control charts test the hypothesis that the process is under the statistical control. However, to realize some hypothesis testing is possible to get a wrong decision. With the control charts is not different, it can be committed two errors, the Type I Error and the Type II Error. The Type I Error supposes that the process is out of statistical control when, actually, it is under statistical control, and the Type II Error supposes that the process is under statistical control when, actually, it is out of statistical control (Ramos *et al.*, 2005). Therefore, it is necessary to control these errors and through the *ARL* statistics it can be evaluated how good is the control chart related to these detected errors, where the *ARL* determine the medium number of necessary samples to obtain a sign that the process is out of statistical control (Type I Error and Type II Error). This Way, this work offer a new methodology to elaborate the *ARL* to the medium control chart, where it is built using an estimator of population standard deviation, based on sample quartiles.

**Keywords:** *ARL*; Quartiles; Control Charts.

## 1 INTRODUÇÃO

A busca das empresas para se tornarem cada vez mais competitivas torna indispensável a utilização dos métodos estatísticos, principalmente o Controle Estatístico da Qualidade (*CEQ*), sendo este utilizado tanto no planejamento industrial, como em técnicas de vendas e para melhoria da qualidade de produtos e/ou serviços. De fato, o uso das técnicas do *CEQ* faz com que melhore a qualidade final do produto. É importante que o produto e/ou serviço seja projetado para atender as expectativas dos consumidores. Portanto, atingir o valor alvo especificado e diminuir a variabilidade, gerando com isso a redução de custos, satisfação dos clientes e valorização do produto (ou serviço) e, conseqüentemente, aumento do lucro, são alguns dos objetivos dos usuários do *CEQ*.

Para verificar se num processo produtivo as variações que estão ocorrendo são decorrentes de causas aleatórias ou de causas especiais de variação, utiliza-se a ferramenta estatística denominada de gráfico de controle (Shewhart, 1931). Um gráfico de controle é composto basicamente por três linhas, a primeira linha é denominada de Limite Superior de Controle (*LSC*), a segunda de Linha Central (*LC*) e a última de Limite Inferior de Controle (*LIC*). Os limites de controle dependem do tamanho amostral e da estatística teste, esta pode ser, por exemplo, a média, o desvio padrão, a mediana, a variância, a amplitude e outras. Dependendo da estatística de teste o gráfico irá avaliar o nível ou a dispersão da variável de interesse, e o nome que o gráfico receberá será gráfico de controle do(a)  $w$  (onde  $w$  é a estatística teste), por exemplo, gráfico de controle da mediana. Segundo Montgomery (2001), de uma forma geral, os limites dos gráficos de controle são construídos da seguinte forma

$$\begin{aligned} LSC &= \mu_w + Z_{(\alpha/2)}\sigma_w \\ LC &= \mu_w \\ LIC &= \mu_w - Z_{(\alpha/2)}\sigma_w, \end{aligned} \quad (1.1)$$

onde  $\mu_w$  e  $\sigma_w$  são, respectivamente, esperança e o erro padrão da variável aleatória  $w$  e  $Z_{(\alpha/2)}$  é o escore associado a distribuição Normal que depende do nível de significância  $\alpha$ , associado ao teste.

Figura 1: Gráfico de Controle de  $w$ .

Para saber se estão ocorrendo causas identificáveis constroem-se os limites de controle a partir da Equação 1.1 e “plotam-se” os valores da estatística teste no gráfico conforme pode ser visto na Figura 1. Se os valores estão entre os limites de controle o processo está sob controle estatístico. Portanto, apenas causas aleatórias estão influenciando na variabilidade do processo, porém, se houver pontos fora destes limites o processo está fora de controle estatístico, ou seja, há causas identificáveis atuando no processo.

**2 DESENVOLVIMENTO**

**2.1 GRÁFICO DE CONTROLE DA MEDIANA**

Para construir o gráfico de controle da mediana (gráfico  $\bar{X}$ ), assume-se que a característica da qualidade  $X$  monitorada, deve ser normalmente distribuída com média  $\mu$  e variância  $\sigma^2$ , portanto,  $E[\bar{X}] = \mu$  e  $Var[\bar{X}] = \sigma^2 \pi^2 / 4n$ . Considerando na Equação 1.1,  $Z_{(\alpha/2)} = 3$ , os usuais limites 3-sigmas, e  $n$  é o tamanho amostral, têm-se que os limites de controle para o gráfico da mediana são obtidos por

$$\begin{aligned}
 LSC &= \mu + 3 \frac{\sigma}{\sqrt{n}} \frac{\pi}{2} \\
 LC &= \mu \\
 LIC &= \mu - 3 \frac{\sigma}{\sqrt{n}} \frac{\pi}{2}.
 \end{aligned}
 \tag{2.1}$$

Quando a média e o desvio padrão populacional não são conhecidos é preciso estimá-los o estimador usual da mediana é o  $\bar{\bar{X}}$ , definido como média das medianas e para estimar o desvio padrão pode-se usar os estimadores  $\bar{R}/d_2$  e  $\bar{S}/c_4$ , onde  $d_2$  e  $c_4$  dependem do tamanho amostral  $n$  (Juran e Gryna, 1992; Mittag e Rinne, 1993).

Ramos (2003) apresenta o estimador  $\overline{IQ}/\xi_n$  e prova que este estimador é não-viesado e consistente para  $\sigma$ . Além disso, Ramos (2003) mostra que  $\overline{IQ}/\xi_n$  é mais resistente (não influenciado por valores extremos) do que  $\bar{R}/d_2$  e  $\bar{S}/c_4$ . Valores de  $\xi_n$  estão tabulados para alguns tamanhos amostrais (Ramos, 2003).

**2.2 RISCO  $\beta$**

Em um teste estatístico têm-se duas hipóteses

$$\begin{cases}
 H_0 : \text{Hipótese básica} \\
 H_1 : \text{Hipótese contrária.}
 \end{cases}
 \tag{2.2}$$

Ao tomar uma decisão utilizando um teste de hipóteses, podem-se cometer dois erros. O Erro Tipo I e o Erro Tipo II, o primeiro é rejeitar a hipótese  $H_0$ , quando  $H_0$  é verdadeira, e o segundo representa não rejeitar  $H_0$ , quando  $H_0$  é falsa. Quando se utiliza um gráfico de controle,  $H_0$  representa a hipótese de que o processo encontra-se sob controle estatístico e  $H_1$  representa a hipótese contrária. Portanto, o Erro Tipo II ou Risco  $\beta$  representa supor que o processo está sob controle estatístico quando, na verdade, encontra-se fora de controle estatístico. O Erro Tipo I é controlado pelo nível de significância ( $\alpha$ ), a partir de  $Z_{(\alpha/2)}$  utilizado na construção dos gráficos de controle. Para se obter o Risco  $\beta$  considera-se, inicialmente, o processo sob controle estatístico. Se apenas o valor do nível do processo muda do estado sob controle estatístico, por exemplo,  $\mu_0$  para outro valor  $\mu_1 = \mu_0 + k\sigma$  fazendo com que o processo fique fora de controle estatístico, onde  $k$  é a variação do nível populacional em função do desvio padrão, então a probabilidade de não detectar esta mudança na primeira amostra subsequente (Risco  $\beta$ ) é

$$\beta = P(LIC \leq w \leq LSC \mid \mu = \mu_1 = \mu_0 + k\sigma), \tag{2.3}$$

onde  $w$  é a estatística teste, neste trabalho  $w$  é a mediana ( $\bar{X}$ ).

A Figura 2 mostra a passagem da distribuição de probabilidade da estatística  $w$  de um processo que está sob controle estatístico (Curva A) para um processo fora de controle estatístico (Curva B). A partir da Figura 2, vê-se também que a mudança no valor do nível do processo de  $\mu_0$  para  $\mu_1 = \mu_0 + k\sigma$  fez com que houvesse maior probabilidade de extrapolação do LSC e, conseqüentemente, fez o número de itens defeituosos do processo aumentar.

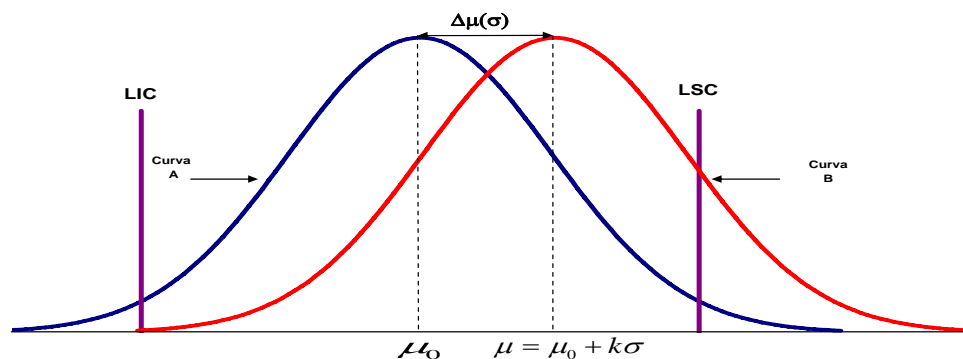


Figura 2: Variação do nível do processo em função de  $k$  desvios padrão.

Como a mediana possui distribuição normal com média  $\mu$  e variância  $\pi\sigma^2/4n$ , pode-se padronizar a Equação 2.3 obtendo-se

$$\beta = \Phi \left[ \frac{LSC - (\mu_0 + k\sigma)}{(\sigma/\sqrt{n})(\pi/2)} \right] - \Phi \left[ \frac{LIC - (\mu_0 + k\sigma)}{(\sigma/\sqrt{n})(\pi/2)} \right], \tag{2.4}$$

simplificando, tem-se.

$$\beta = \Phi \left( Z_{(\alpha/2)} - (2/\pi)k\sqrt{n} \right) - \Phi \left( -Z_{(\alpha/2)} - (2/\pi)k\sqrt{n} \right), \tag{2.5}$$

onde  $\Phi$  denota a distribuição acumulada da normal padrão.

Para ilustrar a utilização da Equação 2.4, supõe-se que inicialmente utiliza-se um gráfico de controle com  $Z_{(\alpha/2)} = 3$ ,  $n = 5$ ,  $\mu = 0$  e  $\sigma = 1$ , objetivando-se determinar a probabilidade de detectar uma mudança na primeira amostra seguinte à mudança do valor do nível do processo. Agora, supõe-se que o nível do processo sofra uma variação de  $k$  desvios padrão, por exemplo,  $k = 1,04\sigma$ , utilizando-se a Equação 2.4 obtém-se

$$\beta = \Phi\left(3 - \frac{1,04 * 2\sqrt{5}}{\pi}\right) - \Phi\left(-3 - \frac{1,04 * 2\sqrt{5}}{\pi}\right) = \Phi(1,52) - \Phi(-4,48) \cong 0,94.$$

Portanto, a probabilidade de que a mudança de  $k = 1,04\sigma$  no nível do processo seja detectada na primeira amostra é somente  $1 - \beta = 0,06$ , ou seja, 6%.

### 3 ARL PARA O GRÁFICO DE CONTROLE DA MEDIANA

ARL representa o número médio de amostras a serem retiradas até se obter um sinal, o qual pode representar um alarme falso ou que o processo desajustou. Durante a utilização do gráfico de controle deseja-se que, em média, o gráfico demore a apresentar alarmes falsos e, em média, detecte rapidamente que o processo está fora de controle estatístico, logo deseja-se que possua um ARL alto para quando o processo está sob controle estatístico e um ARL pequeno quando há alguma causa identificável atuando no processo.

O Risco  $\beta$  representa a probabilidade de não detectar na primeira amostra subsequente que houve uma mudança no nível do processo quando na verdade houve variação, logo a probabilidade de se detectar que o processo está fora de controle estatístico é  $1 - \beta$ , e a probabilidade de que a mudança seja detectada na segunda amostra é  $\beta(1 - \beta)$ ; a probabilidade de que ela seja detectada na terceira amostra é  $\beta^2(1 - \beta)$ , com isso pode-se perceber que se trata de uma distribuição geométrica com o parâmetro da distribuição igual a  $(1 - \beta)$ , com função de probabilidade

$$P(X = x) = (1 - \beta)\beta^{x-1}, \tag{3.1}$$

logo, espera-se que o primeiro ponto fora do limite de controle seja

$$E(x) = \sum_{x=1}^{\infty} x(1 - \beta)\beta^{x-1} = \frac{1}{1 - \beta}. \tag{3.2}$$

A essa esperança dá-se o nome de ARL (*Average Run Length* ou Comprimento Médio de Corrida). Também define-se o ARL como

$$ARL = \frac{1}{\alpha}, \tag{3.3}$$

onde  $\alpha = 1 - \beta$ . Pode-se interpretar  $\alpha$  como a probabilidade de um ponto ser “plotado”, fora dos limites de controle. Portanto, para o exemplo mostrado na seção anterior, obtém-se

$$ARL = \frac{1}{1 - 0,94} = \frac{1}{0,06} \cong 17.$$

Analisando a Figura 2, observa-se que o Risco  $\alpha$  é calculado a partir da probabilidade de um ponto está “plotado” dentro dos limites de controle, porém quando não há uma variação no nível esta probabilidade representa a probabilidade de estar entre os limites de controle, portanto o complementar desta probabilidade representa o  $\alpha$ , que neste caso representa o nível de significância do teste, logo o ARL será calculado, pela Equação 3.3 utilizando-se este  $\alpha$ .

Para construir os ARL's para o Gráfico de Controle da Mediana, inicialmente, monta-se no eixo das abscissas a variação do nível em função do desvio e no eixo das ordenadas os valores dos ARL's. Calcula-se os ARL's através das Equações 3.2 e 3.3, “plotam-se” os ARL's sobre a magnitude da mudança  $\Delta\mu(\sigma)$  e, finalmente, ligam-se os pontos e desta forma está montada a curva do ARL.

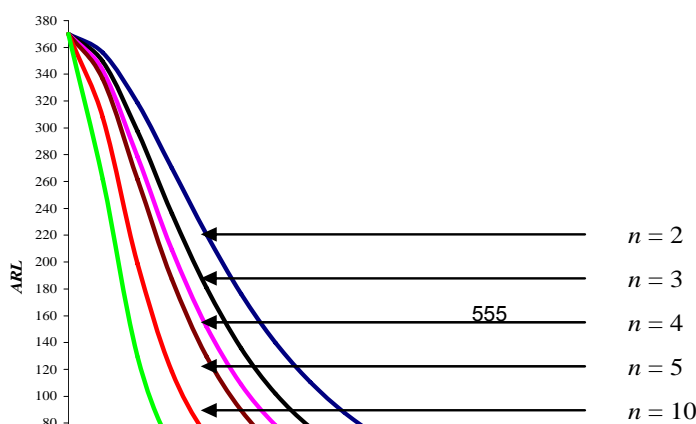


Figura 3: *ARL* para o Gráfico de Controle da Mediana variando o tamanho amostral.

A partir da Figura 3, é possível observar que o Gráfico da Mediana detecta relativamente rápido estas variações, por exemplo, para  $\Delta\mu(\sigma) = 0,4$  e  $n = 20$  o gráfico de controle deve detectar a variação na 35ª amostra (*ARL*).

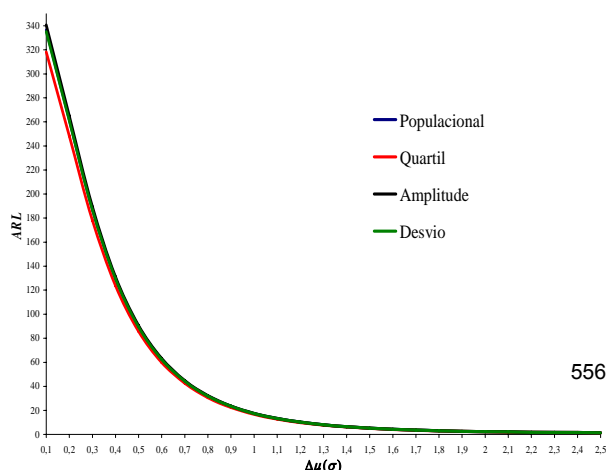
### 3.1 CONSTRUÇÃO DO *ARL* PARA O GRÁFICO DE CONTROLE DA MEDIANA

A metodologia utilizada para construir as curvas do *ARL* para o gráfico de controle da mediana, foi gerar 5000 amostras com a distribuição normal padrão, para tamanhos amostrais  $n = 5, 10, 15$  e  $20$ ; e em seguida foi obtido o *ARL* apenas para variações maiores do que  $0,1\sigma$ , devido ao fato, de para variações menores que  $0,1\sigma$ , a diferença entre o *ARL* populacional e o estimado é muito elevada mesmo que os riscos  $\beta$  (populacional e estimados) sejam próximos, por exemplo, se o Risco  $\beta$  populacional for  $0,9973$  e se os estimados forem de  $0,9972$  ou  $0,9974$  o *ARL* populacional será  $370,37$  e os estimados serão de  $357,14$  e  $384,62$ , respectivamente.

Utiliza-se a média das medianas amostrais  $\bar{\bar{X}}$  para estimar  $\mu$ . Para estimar  $\sigma$ , utilizou-se às estatísticas  $\bar{IQ}/\xi_n$ ,  $\bar{R}/d_2$  e  $\bar{S}/c_4$  baseadas nos quartis amostrais, na amplitude amostral e no desvio padrão amostral, respectivamente.

A partir dos resultados das simulações, obtiveram-se os limites de controle para o gráfico da mediana. Os limites de controle superior e inferior são utilizados para obtenção do Risco  $\beta$ . Os valores do Risco  $\beta$  são então utilizados para obtenção do correspondente valor do *ARL*.

As Figuras 4, 5, 6 e 7 apresentam as curvas do *ARL* para o Gráfico da Mediana, para os tamanhos amostrais  $n = 5; 10; 15$  e  $20$ . As Figuras mostram, para cada um dos tamanhos amostrais  $n$ : (1) o *ARL* populacional, utilizando o Risco  $\beta$  obtido a partir da Equação 2.5, representado pela “Curva Populacional”; (2) o *ARL* quando  $\mu$  é estimada por  $\bar{\bar{X}}$  e  $\sigma$  é estimado por  $\bar{IQ}/\xi_n$ , representado pela “Curva Quartil”; (3) o *ARL* quando  $\mu$  é estimada por  $\bar{\bar{X}}$  e  $\sigma$  é estimado por  $\bar{R}/d_2$ , representado pela “Curva Amplitude” e (4) o *ARL* quando  $\mu$  é estimada por  $\bar{\bar{X}}$  e  $\sigma$  é estimado por  $\bar{S}/c_4$ , representado pela “Curva Desvio”.



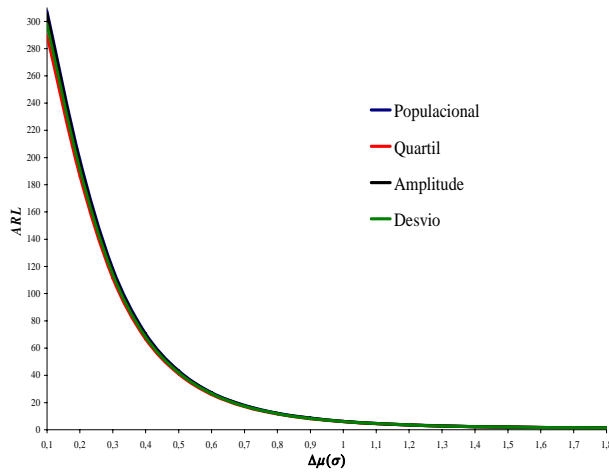


Figura 6:  $ARL$  para o gráfico  $\tilde{X}$  quando  $n = 15$ .  
 Figura 7:  $ARL$  para o gráfico  $\tilde{X}$  quando  $n = 20$ .

Observa-se que para todos os tamanhos amostrais as Curvas dos  $ARL$ 's estimados estão muito próximas da verdadeira curva do  $ARL$  para o Gráfico da Mediana. De fato, os  $ARL$ 's produzidos a partir do estimador  $\bar{X}$  para a média populacional  $\mu$ , e dos estimadores  $\overline{IQ}/\xi_n$ ,  $\bar{R}/d_2$  e  $\bar{S}/c_4$  para estimar o desvio padrão populacional  $\sigma$ , são excelentes estimativas dos verdadeiros  $ARL$ 's para o Gráfico da Mediana.

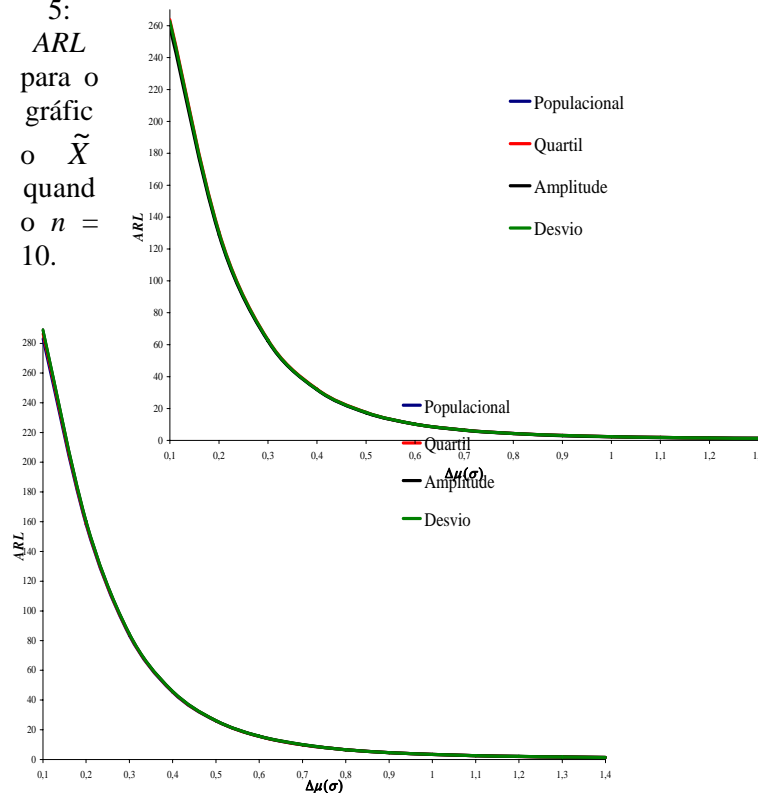
#### 4 CONCLUSÃO

Conclui-se, então, que o  $ARL$  é uma importante ferramenta que associada ao gráfico de controle determina o número médio de amostras até a ocorrência do erro tipo I ou tipo II. Portanto, utilizando-se o  $ARL$  pode-se escolher o melhor tamanho amostral se adapta a produção e a fatores financeiros sem correr risco com relação ao controle de processos. Ainda, vê-se que grandes tamanhos amostrais detectam mais rápido pequenas alterações no nível do processo, e que grandes alterações no nível do processo são detectadas rapidamente, tanto para pequenas amostras quanto para grandes amostras. Finalmente, verifica-se que o estimador baseado nos quartis amostrais é tão bom quanto os usuais estimadores para o desvio padrão para estimar o  $ARL$  e como Ramos (2003) prova que os limites do gráfico de controle produzidos através dos quartis amostrais serão mais consistentes, logo ele deve ser preferível para a construção dos gráficos de controle, pois deseja-se que os limites de controle não sejam influenciados por valores extremos.

#### 5 REFERÊNCIAS BIBLIOGRÁFICAS

[1] JURAN, J.M. e GRAYNA, F.M. *Controle de Qualidade Handbook*, v. 4, São Paulo: Markron Books, 1992.  
 [2] MITTAG, H. J. e RINNE, H. *Statistical Methods of Quality Assurance*, London: Chapman e Hall, 1993.  
 [3] MONTGOMERY, Douglas C. *Introduction to Statistical Quality Control*, New York: John Wiley e Sons, 4<sup>th</sup>. ed., 2001.

Figura 4:  $ARL$  para o gráfico  $\tilde{X}$  quando  $n = 5$ .  
 Figura 5:  $ARL$  para o gráfico  $\tilde{X}$  quando  $n = 10$ .



- [4] RAMOS, Edson. M. L. S. Aperfeiçoamento e Desenvolvimento de Ferramentas do Controle Estatístico da Qualidade - Utilizando Quartis para Estimar o Desvio Padrão. (2003). Tese (Doutorado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis.
- [5] RAMOS, Edson M. L. S.; ALMEIDA, Silvia dos S.; KONRATH, Andréa C.; RIBEIRO, Jaciane do C; MARTINS, Felipe. *Curvas características de operação para o gráfico de controle da média – utilizando um estimador baseado nos quartis para estimar o desvio padrão*. Tecno-lógica, vol. 9, n. 2, p. 41-53, 2005.
- [6] SHEWHART, W. A. *Economic Control of Quality of Manufactured Produc.* Princeton: Van Nostrand Reinhold, 1931.